An Investigative Analysis of Climate Change Using Historical and Modern Weather Data

Hugh Han, Richard Kurtz

Commack High School, Commack, NY

Summary

The average air temperature has been increasing in recent years. The purpose of this study was to compare meteorological differences between the 1800s and 2000s using recorded historical and modern weather data from four different time periods, each separated by approximately 60 years. Historical weather data can be used to observe climate change in recent years and in the years dating back to the early 1800s to note if there has been an acceleration of the rate of temperature change over the past two centuries. In this study, air temperatures and various aspects of precipitation were observed and compared among the four time periods. Historical weather data from 1826-1836 were obtained from a journal recorded at Erasmus Hall High School (Brooklyn, NY), digitized, and compared to data recorded during 1894-1898, 1940-1948, and 2003-2012 from various weather stations in proximity to Erasmus Hall High School. Results showed that annual mean temperatures have increased at higher rates in more recent years and that the number of extreme days each year, defined as the days with temperatures ≥ 90ºF (32.2°C) or temperatures ≤ 10ºF (-12.2°C), increased over time. The number of days with precipitation each year has also increased; however, the amount of annual precipitation has remained constant throughout the 35 years used in this study.

Introduction

The purpose of this study was to use historical weather data to analyze climate change. The rate at which the climate is currently changing is a threat to the environment and the dangers of global warming concern scientists, policy makers, and the general public throughout the world. Rising air temperatures have been correlated to rising ocean temperatures within the last 45 years (1). In previous studies, it was discovered that air temperatures and humidity—possibly related to precipitation—have increased within the last 40 years (2, 3). Likewise, the use of anthropogenic air pollutants such as pesticides and greenhouse gases, which increase the rate of chemical degradation of the ozone layer, have increased as well (4, 5). Correlations have been found between these increasing amounts of pollution and increasing air temperatures (4). It is estimated that between now and 2030, there will be a rate of global temperature change equal to +0.4ºC/decade (6), and that by the year 2100, the air temperature during every season will be roughly 8ºC higher than it is now (7).

Scientists are able to predict these changes and develop future predictions and models by analyzing weather data in the 1900s and 2000s. However, these studies often do not take into account data recorded before then. Due to a lack of scientific or even standardized weather stations before the 1900s, many scientists often rely on weather reconstruction or other indirect methods used to estimate temperatures in the past (1). The nonexistent temperature data is "reconstructed" by using proxy sources such as tree rings, ice cores, or ocean sediments to determine relative air temperatures of a specific location over a given time interval (8). However, these methods may not be as useful as using primary sources of data when determining temperature differences over shorter periods of time such as 200 years (9). Because weather reconstruction via proxies involves a fair amount of possible error, primary sources of weather data such as journals recorded by schools, institutions, or meteorologists may be more reliable than reconstructed data for generalizing trends over a shorter period of time. Although many records of primary weather data exist, they have not yet been utilized to their full potential—more than 60 archives of historical data from the early 1800s exist in New York State alone (10). The analysis of data prior to 1900 will not only allow climate models of a greater time interval to be created, but will also contribute to the development of more accurate representations of future climate change. It is already known that within the last few decades, the temperature has risen significantly (2) and will most likely continue to rise in the future (6, 7). This study takes a novel approach in determining temperature and precipitation changes before the 1900s in the New York City area by utilizing publicly available historical data.

The scientific goal of this study was to determine the rates of climate change using old weather data recorded during the 1800s, 1900s, and 2000s in the New York City area. Analyzing data using a span of ~180 years provides a broad range of time, which allows one to note changes in weather such as the rate of temperature change. Due to the beginning of heavy industrialization, urbanization, and the emission of more carbon dioxide and pollutants, it is likely that the rate of temperature change began to increase significantly during the mid or late 1800s (11). Using historical weather data broadens the analytical perspective of climate, which improves the development of models for the better understanding of climate change. The method formulated in this study may be implemented to analyze temperature changes at any given location as long as there is a source of historical and modern weather data. This method may be useful in the future for creating better climate representations and simulations.

Results

This study was done using temperature and precipitation data recorded at sites in New York City and New Jersey during four separate time frames. Historical weather data from the Erasmus Hall Journal (EHJ) were photographed (Figure 1) and digitized in Microsoft Excel 2010 (Figure 2) before analysis. More modern data were already in digital form and only had to be transposed into Microsoft Excel 2010

(Figure 3). After the data processing, it was confirmed that both the daily minimum temperatures and daily maximum temperatures in almost every month have increased over time (Table 1, Figures 4, 5). It was also found that annual mean temperatures have not only increased, but are increasing at a greater rate than they have in the past (Figure 6). The rate of temperature change over time, or temperature acceleration, has increased drastically and resembles an exponential function. The annual mean temperature increased by 4.6% from 1940-1948 to 2003-2012 (+2.5°F/1.4°C) and was almost triple that of the 1.8% increase from 1894-1898 to 1940-1948 (+0.9°F/0.5°C), which was almost double that of the 0.90% increase from 1826-1836 to 1894-1898 (+0.5°F/0.3°C). The rate of temperature increase in the last ~60 years, from 1940-1948 to 2003-2012, was approximately five times the rate between the ~60 year interval from 1826-1836 and 1894-1898. Also in the last ~60 years, the number of days with maximum temperatures in the interval 90°F ≤ T < 100°F has doubled from ~7 days/year to ~14 days/year (Figure 7). Of the 31 years of data analyzed, the only days with maximum temperatures equal to or over 100°F occurred in the 2003-2012 range (Figure 7). Although the number of extremely hot days (T ≤ 90°F) has increased significantly over time, the number of extremely cold days (T ≥ 10°F) has decreased (Figure 7).

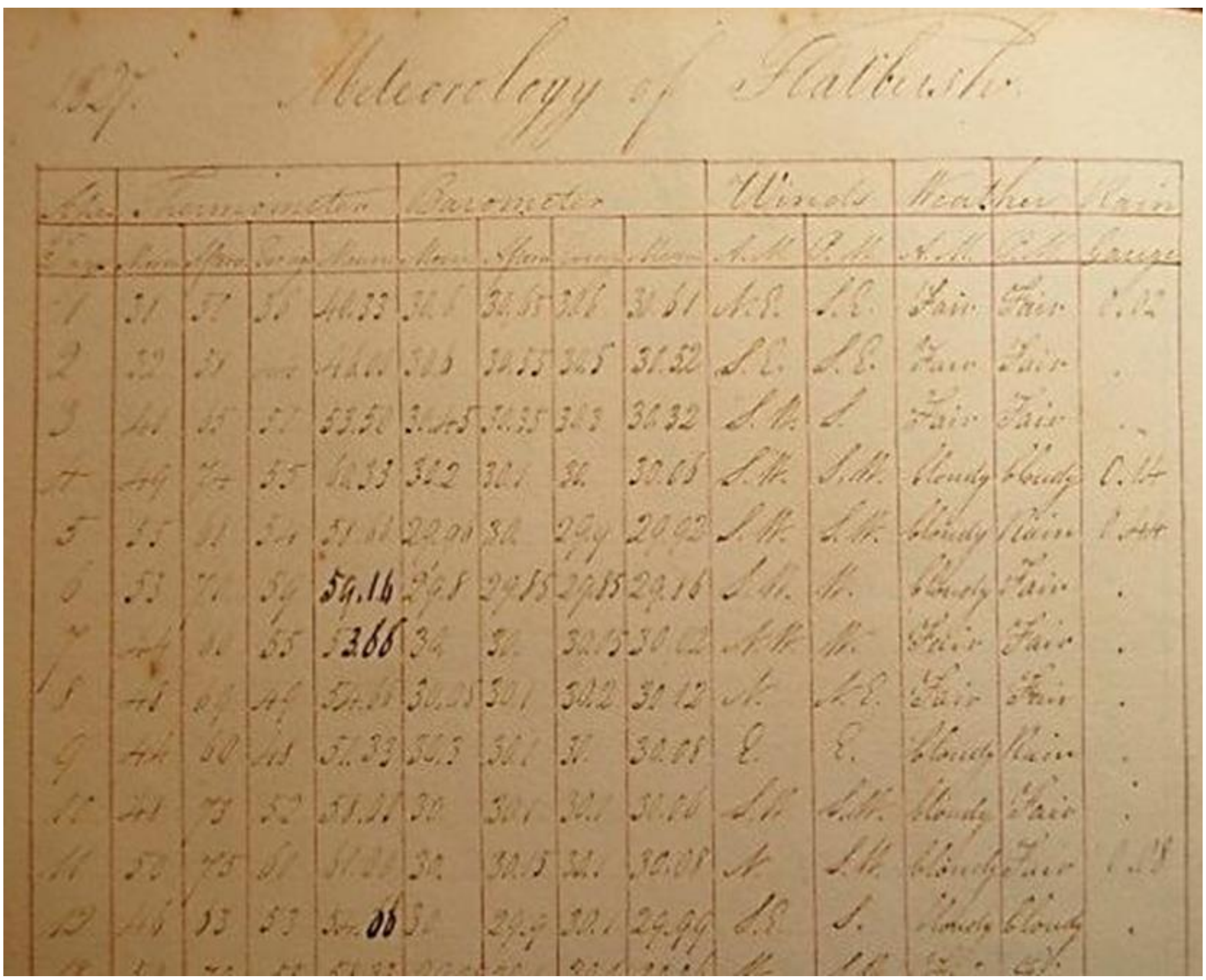
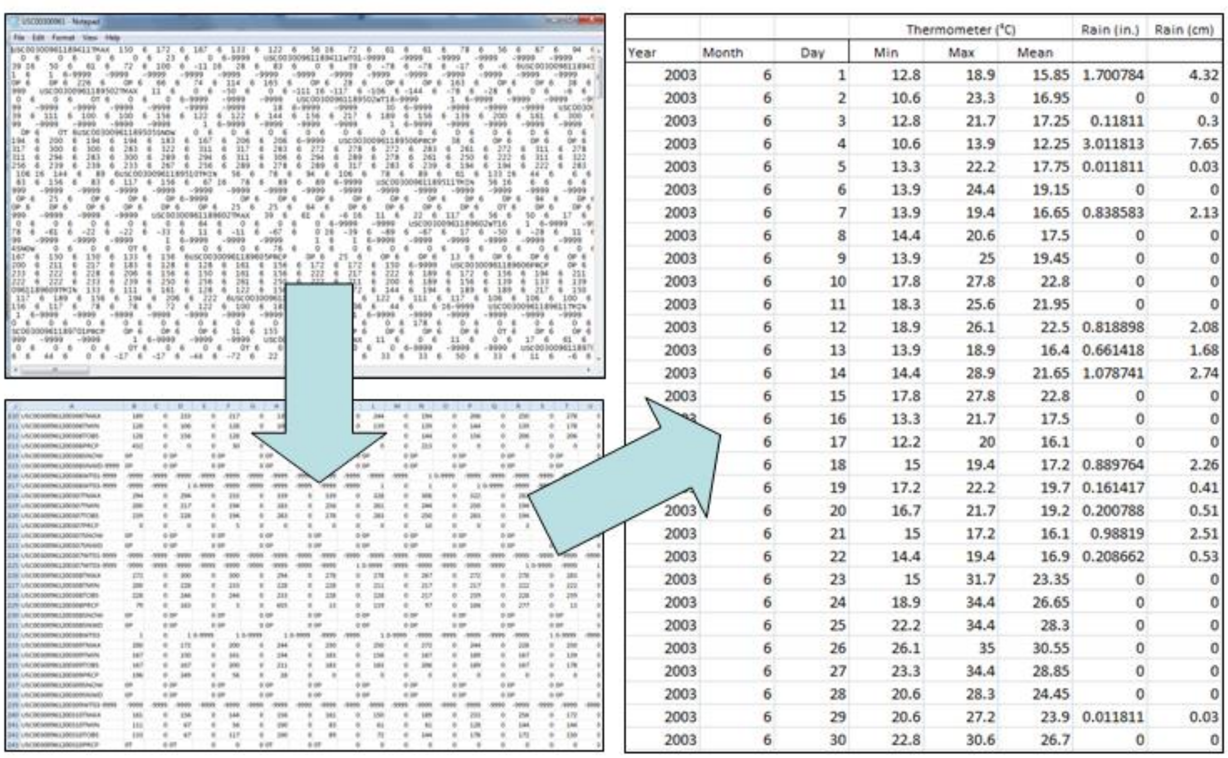


Figure 1: Original data recorded in the Erasmus Hall Journal. This segment of a page regards data recorded on the first 12 days of April 1827.

			Thermometer (ºF)						Rain (in.)	Rain (cm)
Year	Mo.	Day	Morn.	After.	Eve.	Min	Max	Mean		
1827	4	1	38	50	36	36	50	41.33	0.02	0.05
1827	4	2	32	58	44	32	58	44.67		
1827	4	3	40	65	51	40	65	52.00		
1827	4	4	49	74	55	49	74	59.33	0.14	0.36
1827	4	5	55	68	54	54	68	59.00	0.44	1.12
1827	4	6	53	70	59	53	70	60.67		
1827	4	7	44	60	55	44	60	53.00		
1827	4	8	48	69	49	48	69	55.33		
1827	4	9	44	60	58	44	60	54.00		
1827	4	10	48	73	52	48	73	57.67		
1827	4	11	50	75	60	50	75	61.67	0.08	0.2
1827	4	12	46	63	53	46	63	54.00		

Figure 2. Data recorded in the Erasmus Hall High School journal digitized onto Microsoft Excel. This spreadsheet corresponds to a segment of a page regarding data recorded on the first 12 days of April 1827.



			Thermometer (°C)			Rain (in.)	Rain (cm)
Year	Month	Day	Min	Max	Mean		
2003	6	1	12.8	18.9	15.85	1.700784	4.32
2003	6	2	10.6	23.3	16.95	0	0
2003	6	3	12.8	21.7	17.25	0.11811	0.3
2003	6	4	10.6	13.9	12.25	3.011813	7.65
2003	6	5	13.3	22.2	17.75	0.011811	0.03
2003	6	6	13.9	24.4	19.15	0	0
2003	6	7	13.9	19.4	16.65	0.838583	2.13
2003	6	8	14.4	20.6	17.5	0	0
2003	6	9	13.9	25	19.45	0	0
2003	6	10	17.8	27.8	22.8	0	0
2003	6	11	18.3	25.6	21.95	0	0
2003	6	12	18.9	26.1	22.5	0.818898	2.08
2003	6	13	13.9	18.9	16.4	0.661418	1.68
2003	6	14	14.4	28.9	21.65	1.078741	2.74
2003	6	15	17.8	27.8	22.8	0	0
[illegible]	6	16	13.3	21.7	17.5	0	0
[illegible]	6	17	12.2	20	16.1	0	0
[illegible]	6	18	15	19.4	17.2	0.889764	2.26
[illegible]	6	19	17.2	22.2	19.7	0.161417	0.41
2003	6	20	16.7	21.7	19.2	0.200788	0.51
2003	6	21	15	17.2	16.1	0.98819	2.51
2003	6	22	14.4	19.4	16.9	0.208662	0.53
2003	6	23	15	31.7	23.35	0	0
2003	6	24	18.9	34.4	26.65	0	0
2003	6	25	22.2	34.4	28.3	0	0
2003	6	26	26.1	35	30.55	0	0
2003	6	27	23.3	34.4	28.85	0	0
2003	6	28	20.6	28.3	24.45	0	0
2003	6	29	20.6	27.2	23.9	0.011811	0.03
2003	6	30	22.8	30.6	26.7	0	0

Figure 3. Data conversion for weather data from 1894-1898, 1940-1948, and 2003-2012. The upper-left window represents the original .dly file format in WordPad. The bottom-left window represents the data after being directly transferred to Microsoft Excel, and the right window represents the data after being deciphered.

	Minimum Temperature		Maximum Temperature	
Time Periods	p-value	Significant?	p-value	Significant?
1826-1836 vs. 1894-1898	0.0443	Yes	0.0602	No
1894-1898 vs. 1940-1948	0.0273	Yes	0.2675	No
1940-1948 vs. 2003-2012	0.0001	Yes	0.0218	Yes
1826-1836 vs. 1940-1948	< 0.0001	Yes	< 0.0001	Yes
1826-1836 vs. 2003-2012	< 0.0001	Yes	< 0.0001	Yes
1894-1898 vs. 2003-2012	< 0.0001	Yes	0.0218	Yes

Table 1. Student T-Tests: Temperature

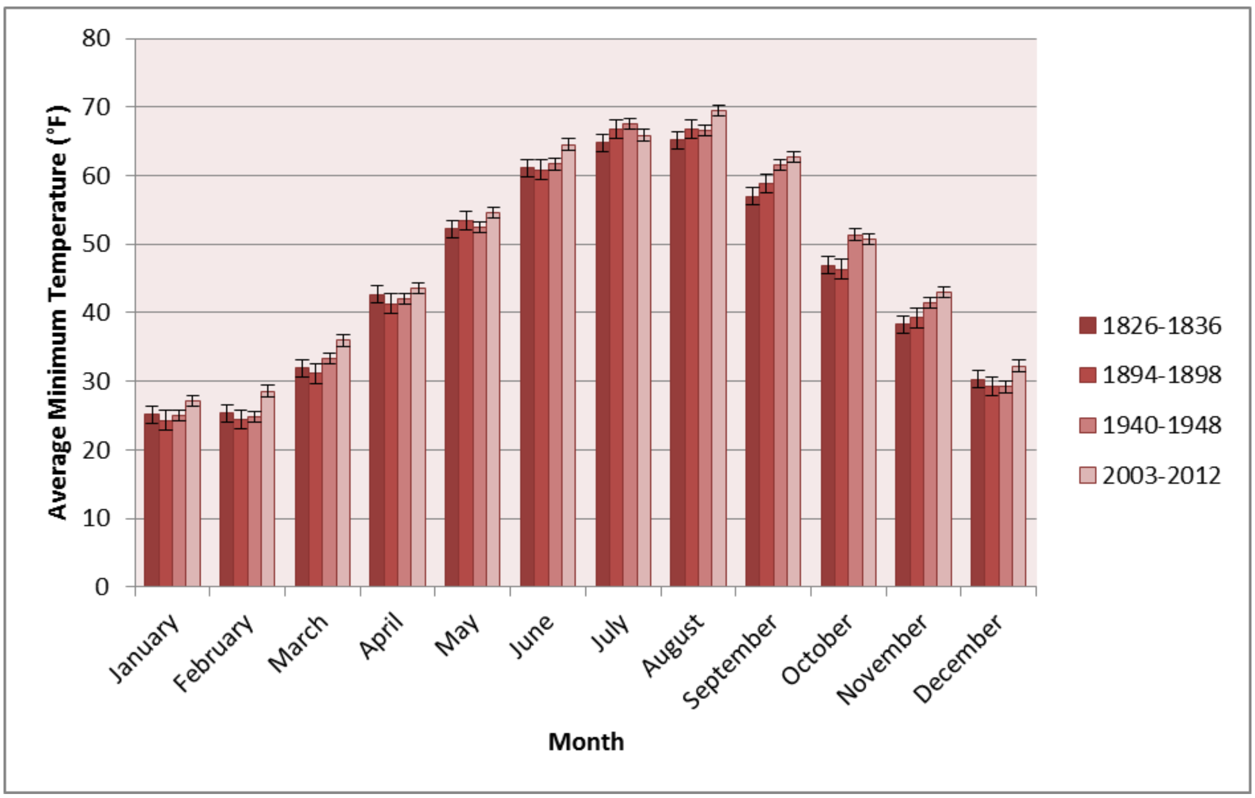


Figure 4. Average Minimum Temperature per Month. Note that between 1826-1836 and 2003-2012, minimum temperatures have increased for every month. Error bars are 99% Confidence Intervals.

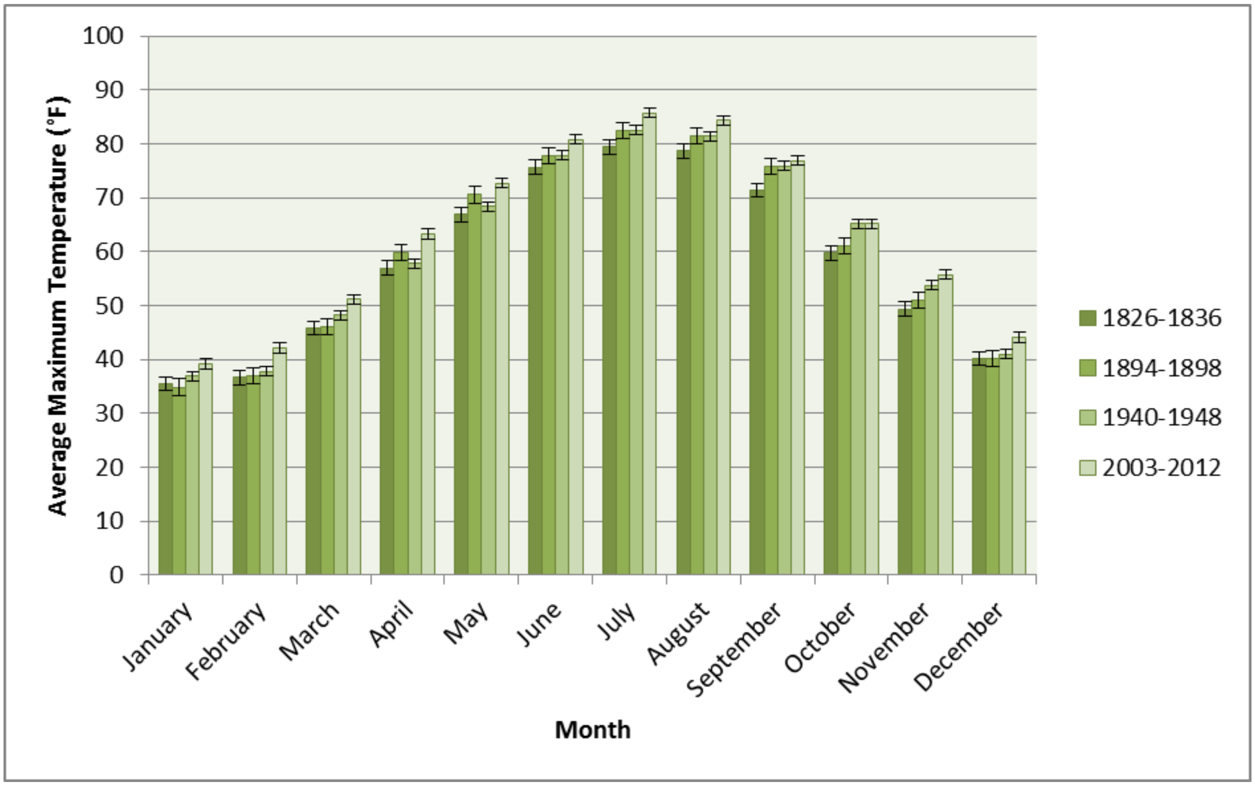


Figure 5. Average Maximum Temperature per Month. Note that between 1826-1836 and 2003-2012, maximum temperatures have increased for every month. Error bars are 99% Confidence Intervals.

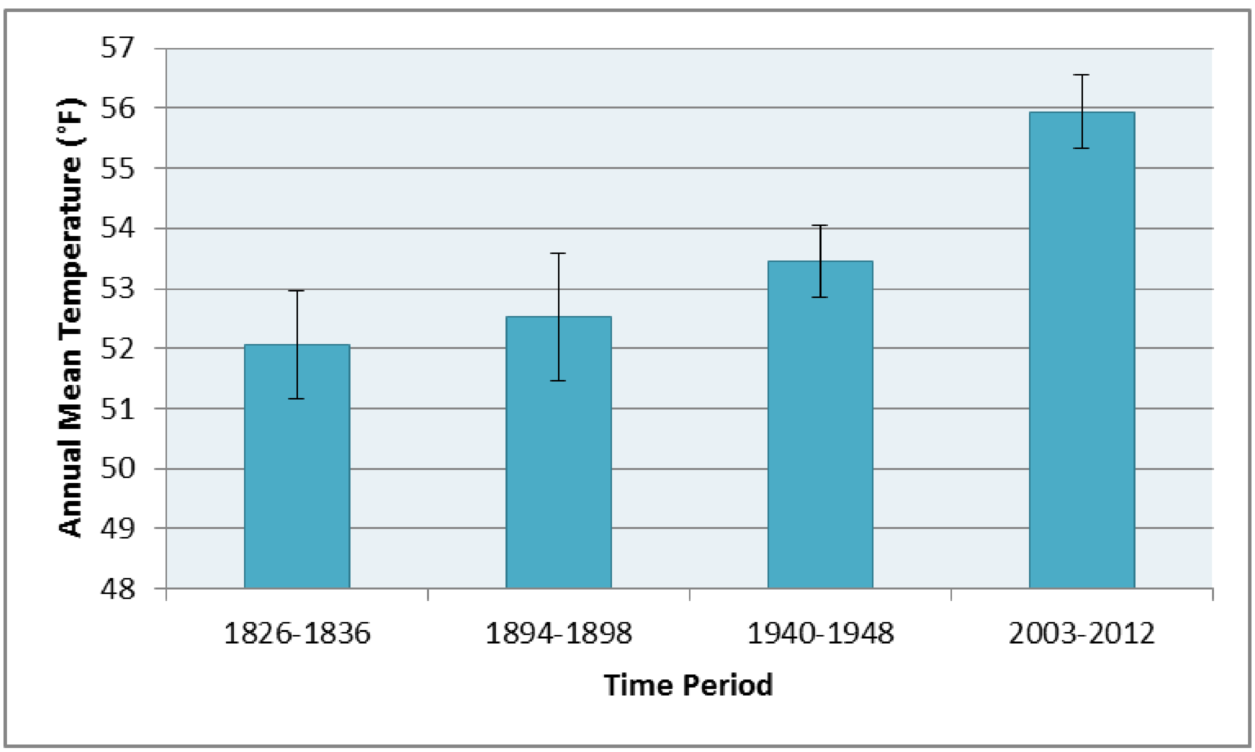


Figure 6. Annual Mean Temperature per Time Period. Note that the annual mean temperature in the New York City region seems to be increasing at an increasing rate. Error bars are 99% Confidence Intervals.

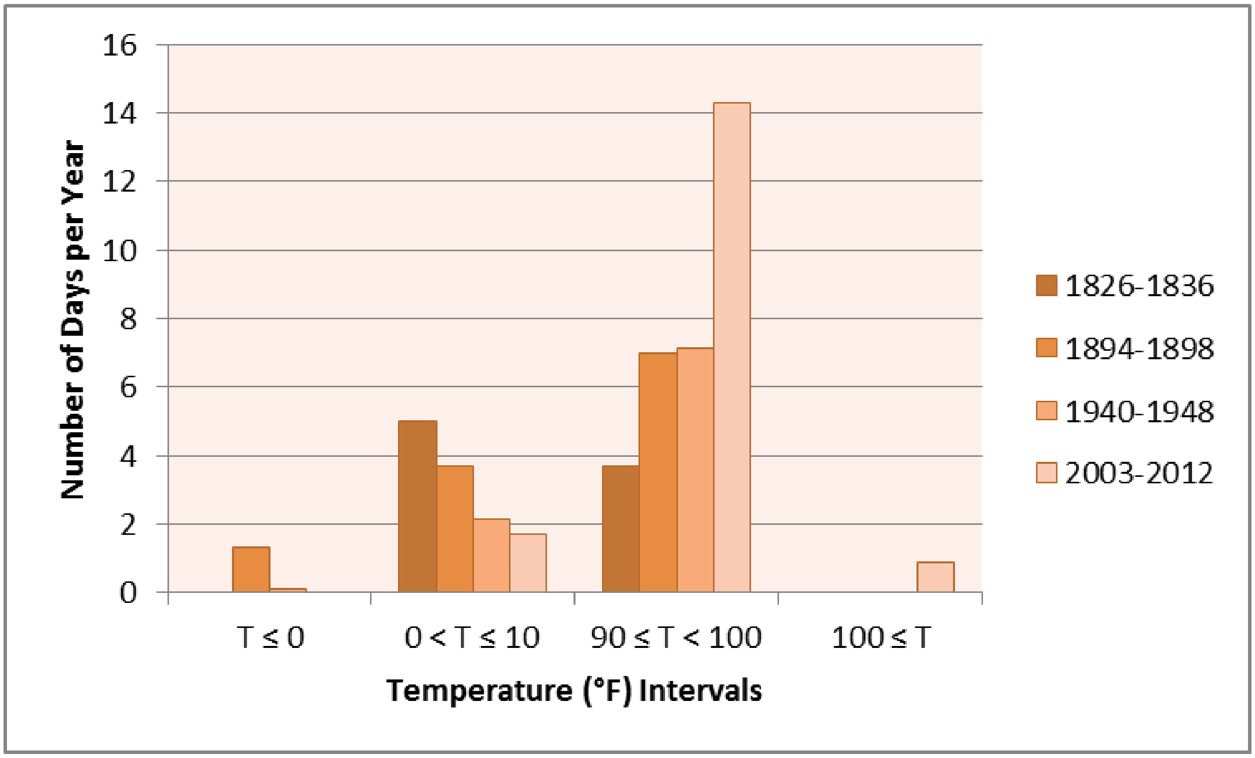


Figure 7. Number of Days per Year vs. Temperature Intervals. In this figure, daily minimum temperatures were used for the colder intervals and daily maximum temperatures were used for the warmer intervals. Note that the number of days lying in the shown “extreme” intervals has increased in more recent years. Error bars are 99% Confidence Intervals.

Interestingly, throughout all four time intervals, there were no significant differences between any of the mean amounts of annual precipitation between time periods except between 1940-1948 and 2003-2012, while the annual frequency of precipitation, defined as the number of days with precipitation each year, has shown a significant difference between each time interval (Table 2, Figures 8, 9). This increase in the frequency of precipitation may be correlated to the increasing temperatures.

	Amount Precipitation		Frequency of Precipitation	
Time Periods	p-value	Significant?	p-value	Significant?
1826-1836 vs. 1894-1898	0.2327	No	< 0.0001	Yes
1894-1898 vs. 1940-1948	0.4288	No	0.0324	Yes
1940-1948 vs. 2003-2012	0.0010	Yes	0.0271	Yes
1826-1836 vs. 1940-1948	0.3631	No	< 0.0001	Yes
1826-1836 vs. 2003-2012	0.2816	No	< 0.0001	Yes
1894-1898 vs. 2003-2012	0.0800	No	0.0001	Yes

Table 2. Student T-Tests: Precipitation

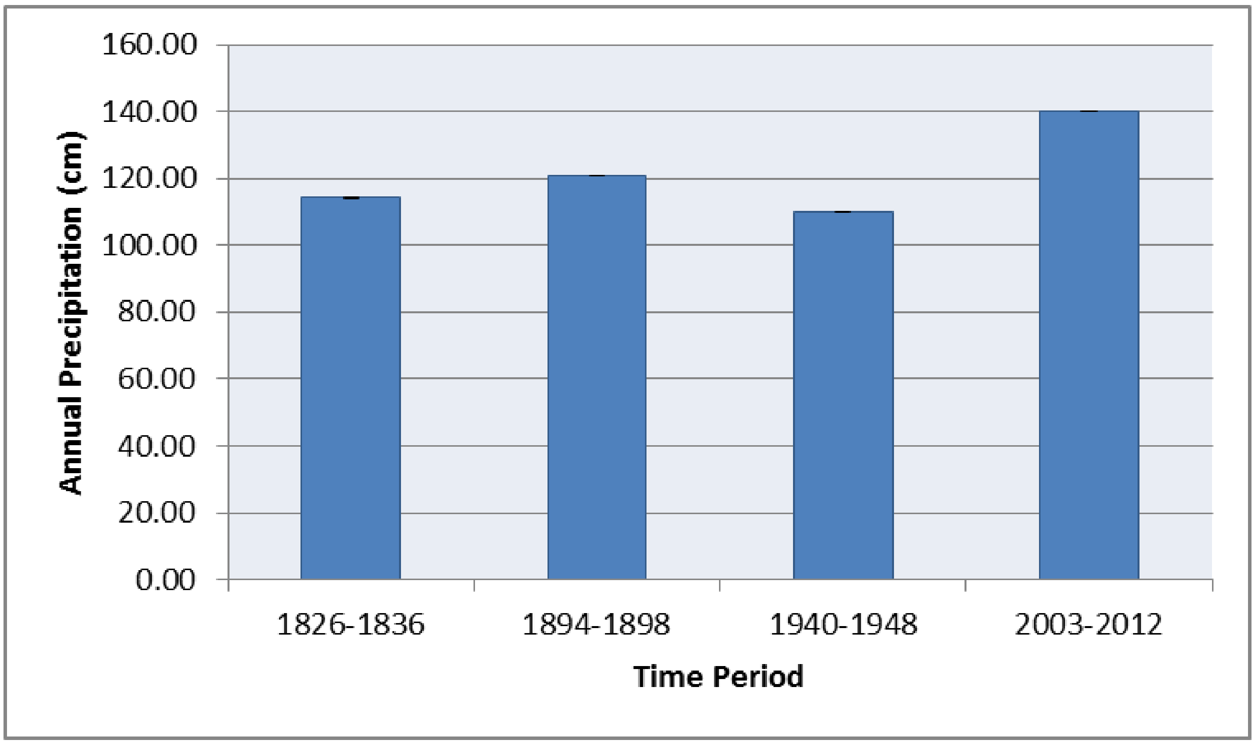


Figure 8. Annual Precipitation during each Time Period. Note that the amount of annual precipitation temperature in the New York City region has not changed significantly. Error bars are 99% Confidence Intervals.

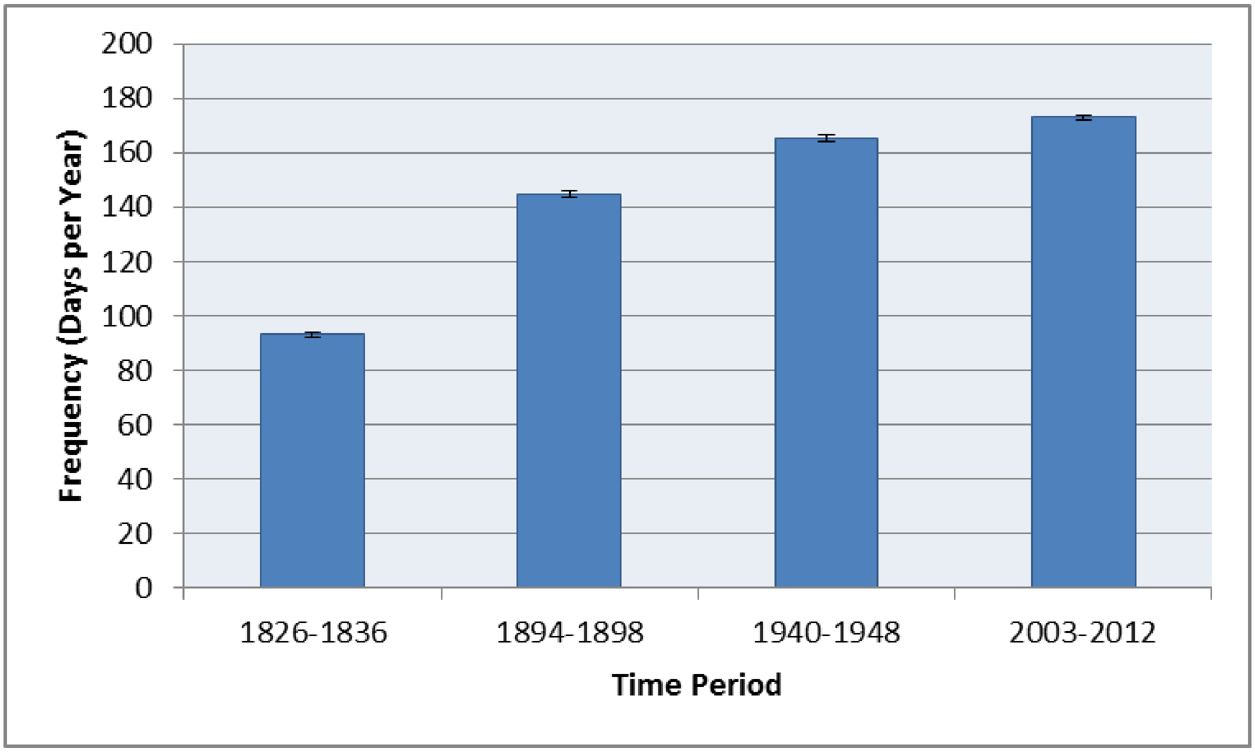


Figure 9. Frequency of Precipitation during each Time Period. Note that the frequency of precipitation temperature in the New York City region has decreased over time. Error bars are 99% Confidence Intervals.

Discussion

In this study using historical weather data as an alternative to weather reconstruction, it was determined that there are currently two aspects of the climate in the New York City region that are increasing: the rate at which air temperatures increase and the frequency of precipitation. The temperature change in the New York City region within the last ~180 years resembles an exponential curve. According to the historical weather data analyzed, this increase can be approximated by the formula f(t) = 52.056(1.0001)t, where t is the time point, starting at t = 1 during the 1826-1836 time point. (The 1894-1898 time point would have a t value of 2, the 1940-1948 a t value of 3, and the 2003-2012 a t value of 4.) f (t) is the annual mean temperature during that time point. Because the data used was collected near a city, there were likely high anthropogenic emissions of carbon dioxide into the atmosphere, causing the exponential temperature increase (11). Since the frequency of precipitation has increased while the annual amount of precipitation has remained roughly constant in the New York City area over the last ~180 years, it is likely that precipitation in the years 2003-2012 occurred more often, but in lesser amounts each time when compared to that in the years 1826-1836. A potential reason for this may be that increasing temperatures are causing a faster rate in the water cycle, causing water to evaporate and then precipitate more frequently (12).

There are many limiting factors to this study due to the lack of specifications about procedures regarding data collection in the EHJ. For example, although the precipitation was recorded every day, the journal does not specify whether it accounted for evaporation, the number of times precipitation data was taken per day, or at what time of day they were recorded. Inconsistent procedures could have led to a misrepresentation of precipitation in the early 1800s. Other limiting factors in this study were the differences in the equipment used to observe data during the different time periods. Data recorded from the 1800s are not as precise as data recorded in the 2000s; data recordings from the 2000s have greater accuracies than even those of the early 1900s. This limitation, however, was accounted for by adjusting temperature recordings by looking at the recorded temperatures at which precipitation was in the form of snow instead of rain (see Materials and Methods). Lastly, although the stations were located within 25 kilometers of Erasmus Hall High School, data could be inconsistent due to elevation differences or varying proximity of bodies of water. A second limitation of this research is the scope of its analysis. Because the data was limited to only those collected in the proximity of New York City, the extent to which temperature and precipitation patterns have changed based on this research may be more representative of metropolitan areas than the entire earth as a whole. However, the viability of this study allows it to be applied to many different locations including more rural areas in the United States and locations in Europe, where there most likely exist primary sources of weather data that dates even further back in time. Furthermore, this study could serve as a baseline for future studies that examine other metropolitan areas in the United States.

A source of error regarding precise data recording was the desire for “pretty” digits found in the EHJ. While digitizing data, it was noted that some units’ digits occurred very frequently while others occurred much less often. Because of the questionable reliability present, a computer program was written to check the frequency of occurrence of units’ digits recorded in the EHJ. This algorithm took in the data from the journal and returned the number of times each units’ digit (0-9) of all values were taken as inputs. Because of the large sample size, it should be expected that each number (0-9) is distributed evenly. However, using a sample size of 4179 recorded values, it was noted that the digit “0” occurred 16.2% more times and the digit “1” 14.6% fewer times than they would have in a uniform distribution (Table 3).

Digit	1	2	3	4	5	6	7	8	9	0
Frequency (n)	357	463	364	443	447	416	394	448	361	486
Frequency (%)	8.54	11.08	8.71	10.60	10.70	9.95	9.43	10.72	8.64	11.63

Table 3. Frequency of Units’ Digit Occurrence in EHJ Temperature Recordings

Despite these limitations, this method of using publicly available historical data to model climate change is an added approach to the investigation of a warming Earth. Using historical data, generalizations on climate change can be made in any area if there is a source of historical and modern data in a sufficient quantity using the method implemented in this study (Table 4). Instead of being further neglected, these viable sources of data can be utilized to their full potentials. In future research, historical weather data dating to the 1700s can be analyzed. For example, future research could include analysis of data recorded in Monticello, VA, where Thomas Jefferson kept historical weather records for the years 1776-1818. Having prior knowledge of hundreds of years of weather data may greatly improve models that represent future climate change, which often utilize less than 100 years of prior data.

Steps	Instructions
1	Obtain sources of historical and modern weather data.
2	Digitize and/or translate data if not already in the form desired.
3	Calculate annual and/or monthly means (and any other aspects desired) for data.
4	Compare these calculations using statistical tests such as a student's t-test.
5	Draw conclusions and generalize trends.

Table 4. Method Used to Create Generalizations Using Historical and Modern Weather Data

Methods

Historical data were obtained from a weather journal (EHJ) from Erasmus Hall High School located in Brooklyn, NY (40.65°N, 73.96°W) and a compilation created by Franklin B. Hough of historical weather data recorded during the 1800s in New York State. Erasmus Hall High School was one of many educational institutions requested by the New York Board of Regents to make meteorological observations (10). The EHJ included data covering the years 1826-1836 recorded by members of the board of Erasmus Hall High School, all but one of whom were either school trustees or principals (10). The EHJ is located in the archives of the New York Historical Society in New York, NY. To obtain the data, pages within the journal including the 10 years' worth of data used in this study were photographed before being manually digitized into Microsoft Excel 2010. Figure 1 depicts a sample photograph taken of a page of the journal regarding the first 12 days of April 1827. Other sources of data from later years were obtained through reports of the weather stations taking meteorological observations within 25 kilometers of Erasmus Hall High School. These sources were weather stations at Bronx, NY (40.85°N, 73.87°W), located approximately 24 kilometers from Erasmus Hall High School, for the years 1894-1898 and 2003-2012 and Sandy Hook Lighthouse, NJ (40.46°N, 74.00°W), located approximately 21 kilometers from Erasmus Hall High School, for the years 1940-1948. Although not optimal, data from different weather stations had to be used, as no one station alone kept archives of data for all of the time frames considered in this study.

The data that were analyzed consisted of air temperatures and daily precipitation. Data used for the 1826-1836 time period consisted of thermometer readings recorded every morning before sunrise, in the afternoon at 3 PM, and in the evening an hour before sunset, to two significant digits. At these specific times, morning and evening temperature recordings were at approximate minimums and afternoon temperature recordings at approximate maximums (10). These observations were taken using thermometers manufactured by John Kendall of New Lebanon Manufactories (10). Measurements of precipitation were taken on a daily basis using a rain gauge that filled with water throughout the day (10). Because the EHJ was a handwritten journal, the data was digitized in Microsoft Excel 2010 before analysis (Figure 2). Unlike the data from the EHJ, both sources from Sandy Hook Lighthouse and Bronx had their daily temperature recordings in the format of "Minimum" and "Maximum" rather than "Morning", "Afternoon", "Evening". However, like the EHJ, precipitation recordings in these two more modern sources of data were taken on a daily basis. Data from these sources were obtained in a .dly file format, deciphered using a key, and transferred into Microsoft Excel 2010 (Figure 3). Aspects of temperature and precipitation from the EHJ were then compared to the three other time intervals: 1894-1898, 1940-1948, and 2003-2012.

In order to compare temperatures from the EHJ to those from modern sources, Microsoft Excel 2010 was used to find the minimum and maximum temperatures of each day using the three given morning, afternoon, and evening values from the EHJ. The minimum temperature was found by taking the lower of the morning and evening temperatures and the maximum temperature was found by taking the afternoon temperature. Next, the average minimum and maximum daily temperatures for each month were compared between the four time periods using unpaired student's t-tests. Then, annual mean temperatures were compared. Afterwards, upper and lower bounds of confidence intervals were generated using the formula $\bar{x} \pm Z\sigma/\sqrt{n}$, where $\bar{x}$ is the mean of the set of values used (not to be confused with mean temperatures), Z is the Z-score (in this case equal to 2.576), σ the standard deviation, and n the sample size. Finally, the number of days each year that had temperatures in the following temperature interval ranges in each time period was determined: $T \leq 0$ºF, 0ºF $< T \leq$ 10ºF, 90ºF $\leq T <$ 100ºF, 100ºF $\leq T$ (in degrees Celsius: $T \leq$ -17.8ºC, -17.8ºC $< T \leq$ -12.2ºC, 32.2ºC $\leq T <$ 37.8ºC, 37.8ºC $\leq T$). Extreme days were defined as the days categorized in the temperature intervals ($T \geq$ 90ºF and $T \leq$ 10ºF). The annual precipitation levels and annual number of days with precipitation during the four time periods were then compared to each other using unpaired student's t-tests. Because the instruments used to measure temperatures and pressures differed in each of the time periods, their accuracies were evaluated before any analysis of data.

The accuracy of the thermometer used during recordings of data in the EHJ was approximated by noting the temperatures at which the precipitation was in the form of either rain or snow in the 29ºF–35ºF (-1.67ºC–1.67ºC) range. It was found that for days with precipitation, 59% of days at 34ºF (1.11ºC), 91% of days above 34ºF, and 14% of days below 34ºF had precipitation in the form of rain. It was estimated that the thermometer used for the EHJ had approximately a +2ºF inaccuracy. Digitized data was then adjusted by subtracting 2ºF from all recordings. The same method was used for data from the other time periods, but no adjustments were made because these data did not display notable inaccuracies.

Acknowledgments

I would like to thank my teacher and research mentor, Mr. Richard Kurtz, who helped me contact the New York Historical Society Museum and Library, where the Erasmus Hall High School Journal was located. He helped me interpret my data and gave me assistance with the direction of my project. He also connected me with Dr. Cary Mock at the University of South Carolina, whom I would also like to thank for helping me locate modern weather stations in the proximity of Erasmus Hall High School. We would like to thank the New-York Historical Society Museum and Library for making available and permitting us to photograph the logbooks containing the meteorological observations from The Regents of the University of the State of New York (1826 to 1850).We would additionally like to thank Dr. David Brooks, President of the Institute for Earth Science Research and Education, for his guidance and contributions throughout the submission process.

References

1. Union of Concerned Scientists. “Past, Present, and Future Temperatures: The Hockeystick FAQ.” Global Warming. Union of Concerned Scientists, 17 Nov. 2005. Web. 1 Oct. 2012. <http://www.ucsusa.org/global_warming/science_and_impacts/science/past-present-and-future.html>.

2. Diaz, Henry F., Thomas W. Giambelluca, and Jon K. Eischeid. “Changes in the Vertical Profiles of Mean Temperature and Humidity in the Hawaiian Islands.” Global and Planetary Change. 77.1-2 (2011): 21-25. Print.

3. Alexander, L. V., et al. “Global Observed Changes in Daily Climate Extremes of Temperature and Precipitation.” Journal of Geophysical Research. 111 (2006): 1984-2012. Print.

4. Noyes, Pamela D., et al. “Toxicology of Climate Change: Environmental Contaminants in a Warming World.” Environment International 35 (2009): 971-86. Print.

5. “Climate Change Indicators in the United States.” United States Environmental Protection Agency. US EPA, 2012. Web. <http://www.epa.gov/climatechange/science/indicators/>.

6. Raes, Frank, and John H. Seinfeld. “New Directions: Climate Change and Air Pollution Abatement: A Bumpy Road.” Atmospheric Environment. 43.32 (2009): 5132-33. Print.

7. Hertig, E., and J. Jacobeit. “Downscaling Future Climate Change: Temperature Scenarios for the Mediterranean Area.” Global and Planetary Change. 63 (2008): 127-31. Print.

8. Von Storch, Hans, et al. “Reconstructing past Climate from Noisy Data.” Science 306 (2004): 679-82. Print.

9. Hough, Franklin B. Results of a Series of Meteorological Observations, Made in Obedience to Instructions from the Regents of the University, at Sundry Academies in the State of New-York, from 1826 to 1850 Inclusive. Albany: Weed, Parsons & Co., Printers, 1855. Print.

10. Blasing, T. J., and Duvick, D. “Reconstruction of precipitation history in North American corn belt using tree rings.” Nature. 307 (1984): 143-145. Print.

11. Falkowski, P., et al. “The global carbon cycle: A test of our knowledge of Earth as a system.” Science 290 (2000): 291-296. Print.

12. Durack, Paul J., Wijffels, Susan E., Matear, Richard J. “Ocean Salinities Reveal Strong Global Water Cycle Intensification During 1950 to 2000.” Science 336 (2012): 455-458. Print.